

Europäisches
PatentamtEuropean
Patent OfficeOffice européen
des brevets

EP04/3244

REC'D 04 AUG 2004

WIPO

PCT

Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03006948.8

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Anmeldung Nr:
Application no.: 03006948.8
Demande no:

Anmeldetag:
Date of filing: 26.03.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Strasse 28
37079 Göttingen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se référer à la description.)

Use of secreted protein for preventing and treating obesity, diabetes and/or
metabolic syndrome

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

A61K38/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT SE SI SK TR LI

26. März 2003

WEICKMANN & WEICKMANN

Patentanwälte
European Patent Attorneys · European Trademark Attorneys

EPO - Munich
73

26. März 2003

DIPL.-ING. H. WEICKMANN (bis 31.1.01)
DIPL.-ING. F. A. WEICKMANN
DIPL.-CHEM. B. HUBER
DR.-ING. H. LISKA
DIPL.-PHYS. DR. J. PRECHTEL
DIPL.-CHEM. DR. B. BÖHM
DIPL.-CHEM. DR. W. WEISS
DIPL.-PHYS. DR. J. TIESMEYER
DIPL.-PHYS. DR. M. HERZOG
DIPL.-PHYS. B. RUTTENSPERGER
DIPL.-PHYS. DR.-ING. V. JORDAN
DIPL.-CHEM. DR. M. DEY
DIPL.-FORSTW. DR. J. LACHNIT

Our Reference:
30275P EP/WWldkt

Applicant:
DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Straße 28

37079 Göttingen
GERMANY

Use of a secreted protein for preventing and treating obesity,
diabetes and/or metabolic syndrome

**Use of a secreted protein for preventing and treating obesity,
diabetes and/or metabolic syndrome**

5

Description

This invention relates to the use of saposin-related proteins, to the use of polynucleotides encoding these, and to the use of effectors thereof in the diagnosis, study, prevention, and treatment of obesity and / or diabetes mellitus and/or metabolic syndrome.

Many human proteins serve as pharmaceutically active compounds. Several classes of human proteins that serve as such active compounds include hormones, cytokines, cell growth factors, and cell differentiation factors. Most proteins that can be used as a pharmaceutically active compound fall within the family of secreted proteins. Secreted proteins are generally produced within cells at rough endoplasmic reticulum, are then exported to the golgi complex, and then move to secretory vesicles or granules, where they are secreted to the exterior of the cell via exocytosis. Examples for commercially used secreted proteins are human insulin, thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, human growth hormone, transforming growth factor beta, tissue plasminogen activator, erythropoietin, and various other proteins. Receptors of secreted proteins, which are membrane-bound proteins, also have potential as therapeutic or diagnostic agents.

It is, therefore, important for developing new pharmaceutical compounds to identify secreted proteins that can be tested for activity in a variety of animal models. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel functions for human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical

diseases, disorders, and/or conditions by using secreted proteins or the genes that encode them.

Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight, frequently resulting in a significant impairment of health. Obesity may be measured by body mass index, an indicator of adiposity or fatness. Further parameters for defining obesity are waist circumferences, skinfold thickness and bioimpedance (see, inter alia, Kopelman (1999), loc. cit.). It is associated with an increased risk for cardiovascular disease, hypertension, diabetes mellitus type II, hyperlipidaemia and an increased mortality rate.

Obesity is influenced by genetic, metabolic, biochemical, psychological, and behavioral factors and can be caused by different reasons such as non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure. As such, it is a complex disorder that must be addressed on several fronts to achieve lasting positive clinical outcome. Since obesity is not to be considered as a single disorder but as a heterogeneous group of conditions with (potential) multiple causes, it is also characterized by elevated fasting plasma insulin and an exaggerated insulin response to oral glucose intake (Koltermann J., (1980) Clin. Invest 65, 1272-1284). A clear involvement of obesity in type 2 diabetes mellitus can be confirmed (Kopelman P.G., (2000) Nature 404, 635-643).

Triglycerides and glycogen are used as the body's fuel energy storage. Glycogen is a large branched polymer of glucose residues that is mainly stored in liver and muscle cells. Glycogen synthesis and degradation is central to the control of the blood glucose level. Triglycerides are stored in

the cytoplasm of adipocytes. Adipocytes are specialized for the synthesis, storage and mobilization of triglycerides. The glycogen and triglyceride metabolism is highly regulated and their interplay is essential for the energy homeostasis of the body. A high glucose level in the adipocytes results in the synthesis of triglycerides as fuel store. A low intracellular glucose level leads to a release of fatty acids, which can be used as substrates for the beta-oxidation to generate energy. Glycogen levels in cells are more variable than triglyceride levels because the turnover of glycogen is higher. Triglycerides are used as long term energy donors once the glycogen stores run low.

Insulin amongst other hormones plays a key role in the regulation of the fuel metabolism. High blood glucose levels stimulate the secretion of insulin by pancreatic beta-cells. Insulin leads to the storage of glycogen and triglycerides and to the synthesis of proteins. The entry of glucose into muscles and adipose cells is stimulated by insulin.

In patients who suffer from diabetes mellitus either the amount of insulin produced by the pancreatic islet cells is too low (diabetes type 1 or insulin dependent diabetes mellitus IDDM) or liver and muscle cells lose their ability to respond to normal blood insulin levels (insulin resistance). In the next stage pancreatic cells become unable to produce sufficient amounts of insulin (diabetes type 2 or non insulin dependent diabetes mellitus NIDDM).

The concept of 'metabolic syndrome' (syndrome X, insulin-resistance syndrome, deadly quartet) was first described in 1966 by Camus and reintroduced in 1988 by Reaven (Camus JP, 1966, Rev Rhum Mal Osteoartic 33(1):10-14; Reaven et al. 1988, Diabetes, 37(12):1595-1607). Today metabolic syndrome is commonly defined as clustering of cardiovascular risk factors like hypertension, abdominal obesity, high blood levels of triglycerides and fasting glucose as well as low blood levels of HDL

cholesterol. Insulin resistance greatly increases the risk of developing the metabolic syndrome (Reaven, 2002, Circulation 106(3): 286-288). The metabolic syndrome often precedes the development of type II diabetes and cardiovascular disease (McCook, 2002, JAMA 288:2709-2716). The control of blood lipid levels and blood glucose levels is the essential for the treatment of the metabolic syndrome (see, for example, Santomauro A. T. et al., (1999) Diabetes, 48(9):1836-1841).

The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin, VCPI, VCPL or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations are not known.

Secreted proteins are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize novel functions for secreted proteins. The present invention advances the state of the art by providing previously unknown functions for human secreted proteins.

Accordingly, the present invention relates to a secreted protein with novel functions in the human metabolism. The present invention discloses specific genes and proteins encoded thereby involved in the regulation of metabolism, especially in the disorders of the metabolic syndrome, obesity, diabetes mellitus, coronary heart disease, eating disorder, cachexia, hypertension, hypercholesterolemia (dyslipidemia), and/or gallstones.

Therefore, the technical problem underlying the present invention was to provide for means and methods for modulating (pathological) metabolic conditions influencing body-weight regulation and/or energy homeostatic

circuits. The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the present invention relates to genes with novel functions in
5 body-weight regulation, energy homeostasis, metabolism, and obesity. The
present invention discloses specific genes involved in the regulation of
body-weight, energy homeostasis, metabolism, and obesity, as well as
related diseases such as diabetes mellitus, eating disorder, cachexia,
hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia),
10 and / or gallstones. The present invention describes saposin-related genes,
their homologous genes and proteins encoded thereby, in particular human
saposin-related genes and proteins (also referred to as prosaposin, variant
Gaucher disease and variant metachromatic leukodystrophy; PSAP; human
pulmonary surfactant-associated protein B (SFTPB); human hypothetical
15 protein FLJ40379), respectively, as being involved in the conditions
mentioned above.

So far, it has not been described that the proteins of the invention and
homologous proteins are involved in the regulation of energy homeostasis
20 and body-weight regulation and related disorders, and thus, no functions in
metabolic diseases and dysfunctions and other diseases as listed above
have been discussed.

In this invention we refer to the proteins encoded by *Drosophila*
25 Saposin-related genes and homologous proteins, preferably human and
murine homologous polypeptides or proteins or sequences encoding these
proteins as saposin-related or as proteins of the invention.

The present invention discloses that saposin-related proteins are regulating
30 the energy homeostasis and fat metabolism especially the metabolism and
storage of triglycerides and glycogen, and polynucleotides, which identify
and encode the proteins disclosed in this invention. The invention also

relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides of the invention. The invention also relates to the use of these polynucleotides, polypeptides and effectors thereof in the diagnosis, study, prevention, and treatment of metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and / or gallstones.

Saposin-related homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids encoding the human saposin-related homologs and the proteins encoded thereby (in particular human prosaposin (PSAP), human pulmonary surfactant-associated protein B (SFTPB), and/or human hypothetical protein FLJ40379).

The glycoprotein Prosaposin (PSAP) is a precursor for four saposins (A,B, C, and D) found in different cellular locations (Morimoto S. et al., (1989) Proc. Nat. Acad. Sci. 86: 3389-3393). Saposins are small lysosomal proteins that serve as activators of various lysosomal lipid-degrading enzymes (Munford R.S. et al., (1995) J. Lipid Res. 36(8): 1653-1663). Prosaposin occurs in various kinds of human secretory fluids such as cerebrospinal fluid, semen, milk, pancreatic juice, and bile. Prosaposin is postulated to mediate neurotrophic signaling events capable of inducing neural differentiation and preventing cell death.

A mutant mouse line in which the sphingolipid activator protein gene has been inactivated by homologous recombination technology showed complex clinical, pathologic, and biochemical abnormalities similar to those of human patients with total saposin deficiency (Fujita N. et al., (1996) Hum Mol Genet. 5(6):711-725.). The main pathology in the brain of affected Sap ^{-/-} mice was hypomyelination and storage of ceramides and

gangliosides. Ceramide storage occurred also in brain, liver, and kidney, and ceramide catabolism is abnormally slow in fibroblasts. Examination of reproductive organs in prosaposin homozygous mutant males shows several abnormalities, indicating that prosaposin is also involved in the development and maintenance of male reproductive organs (Morales C.R. et al., (2000) J Androl 21(6):765-775)

Pulmonary surfactant is a lipid-rich material that prevents lung collapse by lowering surface tension at the air-liquid interface in the alveoli of lung. It is composed primarily of phospholipids, cholesterol and proteins, including four surfactant associated proteins, two collagenous, carbohydrate-binding glycoproteins (PSP-A and PSP-D), and two small hydrophobic proteins (PSP-B and PSP-C). Pulmonary surfactant proteins are involved in respiration, and promote alveolar stability by lowering the surface tension at the air-liquid interface in the peripheral air spaces. PSP-B is a clinically important, developmentally regulated gene. Deficiency of PSPB B was demonstrated in congenital alveolar proteinosis (Nogee et al., (1993) N Engl J Med. 328(6):406-410).

The invention particularly relates to a nucleic acid molecule encoding a polypeptide contributing to regulating the energy homeostasis and the metabolism of triglycerides and glycogen, wherein said nucleic acid molecule comprises

- (a) the nucleotide sequence of the *Drosophila* saposin-related gene or a mammalian, e.g. human saposin-related homolog (in particular human PSAP, SFTPB and/or FLJ40379), and/or a sequence complementary thereto,
- (b) a nucleotide sequence which hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a sequence of (a),
- (c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,

- 5 (d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the amino acid sequences of a saposin-related protein, preferably a mammalian, e.g. human saposin-related homolog (in particular PSAP, SFTPB, and/or FLJ40379),
- 10 (e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or
- (f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of at least 15 bases, preferably at least 20 bases, more preferably at least 25 bases and most preferably at least 50 bases.

15 The present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity, fragments of said genes, polypeptides encoded by said genes or fragments thereof, and effectors e.g. antibodies, biologically active nucleic acids, such as

20 antisense molecules, RNAi molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

25 The ability to manipulate and screen the genomes of model organisms such as the fly *Drosophila melanogaster* provides a powerful tool to analyze biological and biochemical processes that have direct relevance to more complex vertebrate organisms due to significant evolutionary conservation of genes, cellular processes, and pathways (see, for example, Adams M. D. et al., (2000) *Science* 287: 2185-2195). Identification of novel gene

30 functions in model organisms can directly contribute to the elucidation of correlative pathways in mammals (humans) and of methods of modulating them. A correlation between a pathology model (such as changes in

triglyceride levels as indication for metabolic syndrome including obesity) and the modified expression of a fly gene can identify the association of the human ortholog with the particular human disease.

5 In one embodiment, a forward genetic screen is performed in fly displaying a mutant phenotype due to misexpression of a known gene (see, Johnston Nat Rev Genet 3: 176-188 (2002); Rorth P., (1996) Proc Natl Acad Sci U S A 93: 12418-12422). In this invention, we have used a genetic screen to identify mutations of Saposin-related homologous genes that cause
10 changes in the body weight, which are reflected by a significant change of triglyceride levels. Additionally glycogen levels are analysed.

One resource for screening was a *Drosophila melanogaster* stock collection of EP-lines. The P-vector of this collection has Gal4-UAS-binding sites
15 fused to a basal promoter that can transcribe adjacent genomic *Drosophila* sequences upon binding of Gal4 to UAS-sites (Brand & Perrimon (1993) Development 118:401-415; Rorth P., supra). This enables the EP-line collection for overexpression of endogenous flanking gene sequences. In addition, without activation of the UAS-sites, integration of the EP-element
20 into the gene is likely to cause a reduction of gene activity, and allows determining its function by evaluating the loss-of-function phenotype.

To isolate genes with a function in energy homeostasis, several thousand EP-lines were tested for their triglyceride/glycogen content after a
25 prolonged feeding period (see Examples for more detail). Lines with significantly changed triglyceride/glycogen content were selected as positive candidates for further analysis. The change of triglyceride/glycogen content due to the loss of a gene function suggests gene activities in energy homeostasis in a dose dependent manner that
30 control the amount of energy stored as triglycerides or glycogen.

In this invention, the content of triglycerides and glycogen of a pool of flies with the same genotype was analyzed after feeding for six days using a triglyceride and a glycogen assay. Male flies homozygous for the integration of vectors for *Drosophila* lines HD-EP(3)36824 were analyzed in assays measuring the triglyceride/glycogen contents of these flies (illustrated in more detail in the Examples section). The results of the triglyceride/glycogen content analysis are shown in Figure 1.

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(3)36824) integration. Using those isolated genomic sequences public databases like Berkeley *Drosophila* Genome Project (GadFly; see also FlyBase (1999) *Nucleic Acids Research* 27:85-88) were screened thereby identifying the integration site of the vectors, and the corresponding gene, described in more detail in the Examples section. The molecular organization of the gene is shown in Figures 2.

The *Drosophila* genes and proteins encoded thereby with functions in the regulation of triglyceride metabolism were further analysed in publicly available sequence databases (see Examples for more detail) and mammalian homologs were identified (see Figure 3).

In addition, we identified prosaposin (PSAP) as secreted factor expressed in developing mouse pancreas, as described in more detail in the Examples section.

The function of the mammalian homologs in energy homeostasis was further validated in this invention by analyzing the expression of the transcripts in different tissues and by analyzing the role in adipocyte differentiation.

Expression profiling studies (see Examples for more detail) confirm the particular relevance of the protein of the invention as regulators of energy

metabolism in mammals. Transcripts of prosaposin are found in several tissues of mammals with high expression levels in brain tissues (hypothalamus), in kidney, heart and spleen. In addition, prosaposin shows high expression in brown adipose tissue (BAT) and white adipose tissue (WAT) (see FIGURE 4A). Brown adipose tissue is a well characterized tissue which is well developed in newborn mammals, including humans. One important task of BAT is to generate heat and maintain body temperature homeostasis in newborn. Thus, an expression of the protein of the invention in adipose tissues is confirming a role in the regulation of energy homeostasis and thermogenesis.

Further, we show that mammalian prosaposin is regulated by fasting and by genetically induced obesity. In this invention, we used mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor) mice) to study the expression of prosaposin. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning et al, 1998, Mol. Cell. 2:449-569). We found, for example, that the expression of prosaposin is strongly upregulated in liver of fasted and ob/ob mice (see Figure 4B). In addition, a marked upregulation can be observed in the metabolically active tissue (for example, white adipose tissue (WAT)) of genetically obese (ob/ob) (see FIGURE 4B).

Susceptible wild type mice (for example C57Bl/6) show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, if fed a high fat diet. In such mice, the most prominent response with regard to metabolically active tissues was observed. In those mice, the expression of prosaposin is significantly enhanced in white adipose tissue (see Figure 4C) and in liver and muscle tissues, supporting a hypothesis that the protein of the invention is a modulator of adipogenesis.

In addition, we show in this invention that the the protein of the invention mRNA is up-regulated during adipocyte differentiation in vitro (see EXAMPLES for more detail), suggesting a role as modulator of adipocyte lipid accumulation. With regard to changes in expression intensity during the differentiation of preadipocytes to adipocytes, a enhancement in relative signal intensity can be observed for the protein of the invention during the in vitro differentiation program of 3T3-L1 (see Figure 4D). Thus, we conclude that prosaposin or variants or processing products thereof have a function in the metabolism of mature adipocytes.

Microarrays are analytical tools routinely used in bioanalysis. A microarray has molecules distributed over, and stably associated with, the surface of a solid support. The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as monitoring gene expression, drug discovery, gene sequencing, gene mapping, bacterial identification, and combinatorial chemistry. One area in particular in which microarrays find use is in gene expression analysis (see Example 4). Array technology can be used to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition, condition, disease, or disorder.

Microarrays may be prepared, used, and analyzed using methods known in the art (see for example, Brennan, T.M. et al. (1995) U.S. Patent No.

5,474,796, Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:21502155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

10 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to
15 identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular,
20 this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

25

As determined by Microarray analysis, prosaposin shows differential expression in human primary adipocytes. A strong downregulation is observed concerning the expression of prosaposin is during the human adipocyte differentiation (see FIGURE 5). The prosaposin protein in
30 preadipocytes has the potential to inhibit adipose differentiation at a very early stage. Therefore, the prosaposin protein might play an essential role in adipogenesis. The results are suggesting a role of prosaposin in the

regulation in human metabolism, for example, as modulator (for example, inhibitor) of adipogenesis. Thus, prosaposin is a strong candidate for a pharmaceutical composition for the treatment of conditions related to human metabolism, such as obesity, diabetes, and/or metabolic syndrome.

5

The invention also encompasses polynucleotides that encode the proteins of the invention and homologous proteins. Accordingly, any nucleic acid sequence, which encodes the amino acid sequences of the proteins of the invention and homologous proteins, can be used to generate recombinant molecules that express the proteins of the invention and homologous proteins. In a particular embodiment, the invention encompasses a nucleic acid encoding *Drosophila* saposin-related, or their mammalian, e.g. human homologs; referred to herein as the proteins of the invention. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

10

15

20

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on the melting temperature (T_m) of the nucleic acid binding complex or probe, as taught in Wahl & Berger (1987: Methods Enzymol. 152:399-407) and Kimmel (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 65°C, particularly for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at

25

30

65°C, a positive hybridization signal is observed. Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides or peptide mimetics having a length of at least 4, preferably at least 6 and up to 50 amino acids.

Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention and homologous proteins. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structures or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with the others, one or more times in a given sequence.

The nucleic acid sequences encoding the proteins of the invention and homologous proteins may be extended utilizing a partial nucleotide

sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements.

5 In order to express a biologically active protein, the nucleotide sequences encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding
10 the proteins and the appropriate transcriptional and translational control elements. Regulatory elements include for example a promoter, an initiation codon, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the Cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin
15 promoter (see, Soria et al., 2000, Diabetes 49:157), SOX2 gene promoter (see Li et al., (1998) Curr. Biol. 8:971-974), Msi-1 promoter (see Sakakibara et al., (1997) J. Neuroscience 17:8300-8312), alpha-cardia myosin heavy chain promoter or human atrial natriuretic factor promoter
20 (Klug et al., (1996) J. clin. Invest 98:216-224; Wu et al., (1989) J. Biol. Chem. 264:6472-6479) or (iii) inducible promoters such as the tetracycline inducible system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro
25 recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

30

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention and

homologous proteins may be ligated to a heterologous sequence to encode a fusion protein.

5 A variety of expression vector/host systems may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include, but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus, adenovirus, 10 adeno-associated virus, lentivirus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.

15 The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization and/or amplification using probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants 20 containing DNA or RNA encoding the corresponding protein. As used herein 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or primer.

25

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide sequences include oligo-labeling, nick translation, 30 end-labeling of labeled RNA probes, PCR amplification using a labeled nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn,

(Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

5 The presence of proteins of the invention in a sample can be determined by immunological methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for determining protein activity are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and
10 fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a
15 Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic
20 agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

The nucleic acids encoding the proteins of the invention can be used to generate transgenic animal or site specific gene modifications in cell lines.
25 Transgenic animals may be made through homologous recombination, where the normal locus of the genes encoding the proteins of the invention is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like. The modified cells or animal
30 are useful in the study of the function and regulation of the proteins of the invention. For example, a series of small deletions and/or substitutions may be made in the genes that encode the proteins of the invention to

determine the role of particular domains of the protein, functions in pancreatic differentiation, functions in adipocyte differentiation, etc.

Specific constructs of interest include anti-sense molecules, which will
5 block the expression of the proteins of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z, may be introduced in the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detected change in phenotype.

10 One may also provide for expression of the genes of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the proteins of the invention in cells in which they are not normally produced,
15 one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA
20 constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and/or negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic
25 cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in presence of leukemia inhibiting factor (LIF).

When ES or embryonic cells or somatic pluripotent stem cells have been
30 transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a

selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected. The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, etc.

Diagnostics and Therapeutics

The data disclosed in this invention show that the nucleic acids and proteins of the invention and effector molecules thereof are useful in diagnostic and therapeutic applications implicated, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and / or gallstones. Hence, diagnostic and therapeutic uses for the proteins of the invention nucleic acids and proteins of the invention are, for example but not limited to, the following: (i) protein therapeutic, (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic

and/or prognostic marker, (v) gene therapy (gene delivery/gene ablation), (vi) research tools, and (vii) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues).

5

The nucleic acids and proteins of the invention and effectors thereof are useful in diagnostic and therapeutic applications implicated in various applications as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, but not limited to, in metabolic disorders as described above.

15

The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that bind immunospecifically to the novel substances of the invention may be used in therapeutic or diagnostic methods.

20

For example, in one aspect, antibodies, which are specific for the proteins of the invention and homologous proteins, may be used directly as an effector, e.g. an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the protein. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

25

30

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. It is preferred that the peptides, fragments or oligopeptides used to induce antibodies to the protein have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

Monoclonal antibodies to the proteins may be prepared using any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Köhler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R. J. et al. (1984) *Proc. Natl. Acad. Sci.* 80:2026-2030; Cole, S. P. et al. (1984) *Mol. Cell Biol.* 62:109-120).

In addition, techniques developed for the production of 'chimeric antibodies', the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) *Proc. Natl. Acad. Sci.* 81:6851-6855; Neuberger, M. S. et al (1984) *Nature* 312:604-608; Takeda, S. et al. (1985) *Nature* 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for the proteins of the invention and homologous proteins. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D. R. (1991) *Proc. Natl. Acad. Sci.* 88:11120-11123). Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels

of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

- 5 Antibody fragments, which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not limited to, the $F(ab')_2$ fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of $F(ab')_2$ fragments.
- 10 Alternatively, Fab expression-libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

- 15 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site,
- 20 monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering protein epitopes are preferred, but a competitive binding assay may also be employed (Maddox, supra).

- 25 In another embodiment of the invention, the polynucleotides or fragments thereof or nucleic acid effector molecules such as antisense molecules, aptamers, RNAi molecules or ribozymes may be used for therapeutic purposes. In one aspect, aptamers, i.e. nucleic acid molecules, which are capable of binding to a protein of the invention and modulating its activity, may be generated by a screening and selection procedure involving the use
- 30 of combinatorial nucleic acid libraries.

In a further aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding the proteins of the invention and homologous proteins. Thus, antisense molecules may be used to modulate protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules complementary to the polynucleotides of the genes encoding the proteins of the invention and homologous proteins. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention and homologous proteins can be turned off by transforming a cell or tissue with expression vectors, which express high levels of polynucleotides that encode the proteins of the invention and homologous proteins or fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, e.g. DNA, RNA or nucleic acid analogues such as PNA, to the control regions of the genes encoding the proteins of the invention and homologous proteins, i.e., the promoters, enhancers, and

introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and + 10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it cause inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J. E. et al. (1994) In; Huber, B. E. and B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples, which may be used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention and homologous proteins. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Nucleic acid effector molecules, e.g. antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for

chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above.

Such pharmaceutical compositions may consist of the nucleic acids and the proteins of the invention and homologous nucleic acids or proteins, antibodies to the proteins of the invention and homologous proteins, mimetics, agonists, antagonists or inhibitors of the proteins of the invention and homologous proteins or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones. The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used to determine the appropriate concentration range and route of

administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active ingredient, for example the nucleic acids or the proteins of the invention or fragments thereof or antibodies, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation. Normal dosage amounts may vary from 0.1 to 100,000 μ g, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly,

delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

5 In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or associated with over- or underexpression of the proteins of the invention and homologous proteins or in assays to monitor patients being treated with the proteins of the invention and homologous proteins, or effectors thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include
10 methods which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules, which are known in the art may be used
15 several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene expression. Normal or standard values for gene
20 expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of protein expressed in control
25 and disease, samples e.g. from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

30 In another embodiment of the invention, the polynucleotides specific for the proteins of the invention and homologous proteins may be used for diagnostic purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs.

The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

In one aspect, hybridization with probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding the proteins of the invention and homologous proteins or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be DNA or RNA and are preferably derived from the nucleotide sequence of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be used for the diagnosis of conditions or diseases, which are associated with the expression of the proteins. Examples of such conditions or diseases include, but are not limited to, metabolic diseases and disorders, including obesity, diabetes, and/or metabolic syndrome. Polynucleotide sequences specific for the proteins of the invention and homologous proteins may also be used to monitor the progress of patients receiving treatment for metabolic diseases and disorders, including obesity, diabetes, and/or metabolic syndrome. The polynucleotide sequences may be used qualitative or quantitative assays, e.g. in Southern or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays

utilizing fluids or tissues from patient biopsies to detect altered gene expression.

5 In a particular aspect, the nucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be useful in assays that detect activation or induction of various metabolic diseases and disorders, including obesity, diabetes, and/or metabolic syndrome. The nucleotide sequences may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of
10 hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. The presence of altered levels of nucleotide sequences encoding the proteins of the invention and homologous proteins in the sample indicates the presence of the associated disease. Such assays may also be used to
15 evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

20 In order to provide a basis for the diagnosis of a disease associated with expression of the proteins of the invention and homologous proteins, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a fragment thereof, which is specific for the nucleic acids encoding the proteins of the
25 invention and homologous nucleic acids, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may
30 be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established

and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to metabolic diseases such as described above the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding the proteins of the invention and homologous proteins may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5prime.fwdarw.3prime) and another with antisense (3prime.rarw.5prime), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

In another embodiment of the invention, the nucleic acid sequences may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH,

FACS or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C. M. (1993) Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154.

5 FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of

10 the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or predisposition to a specific disease, may help to delimit the region of DNA associated with that genetic disease.

The nucleotide sequences of the subject invention may be used to detect

15 differences in gene sequences between normal, carrier or affected individuals. An analysis of polymorphisms, e.g. single nucleotide polymorphisms may be carried out. Further, in situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for

20 extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms or parts thereof, by physical mapping. This provides

25 valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent

30 associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences

in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

5 In another embodiment of the invention, the proteins of the invention, their catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds in any of a variety of drug screening techniques. One can identify effectors, e.g. receptors, enzymes, proteins, ligands, or substrates that bind to, modulate or mimic the action of one or more of the
10 proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between the proteins of the invention and the agent tested, may be measured. Agents could also, either directly or indirectly, influence the
15 activity of the proteins of the invention.

In addition activity of the proteins of the invention against their physiological substrate(s) or derivatives thereof could be measured in cell-based assays. Agents may also interfere with posttranslational
20 modifications of the protein, such as phosphorylation and dephosphorylation, farnesylation, palmitoylation, acetylation, alkylation, ubiquitination, proteolytic processing, subcellular localization and degradation. Moreover, agents could influence the dimerization or oligomerization of the proteins of the invention or, in a heterologous
25 manner, of the proteins of the invention with other proteins, for example, but not exclusively, docking proteins, enzymes, receptors, or translation factors. Agents could also act on the physical interaction of the proteins of this invention with other proteins, which are required for protein function, for example, but not exclusively, their downstream signaling.

30

Methods for determining protein-protein interaction are well known in the art. For example binding of a fluorescently labeled peptide derived from the

interacting protein to the protein of the invention, or vice versa, could be detected by a change in polarisation. In case that both binding partners, which can be either the full length proteins as well as one binding partner as the full length protein and the other just represented as a peptide are fluorescently labeled, binding could be detected by fluorescence energy transfer (FRET) from one fluorophore to the other. In addition, a variety of commercially available assay principles suitable for detection of protein-protein interaction are well known in the art, for example but not exclusively AlphaScreen (PerkinElmer) or Scintillation Proximity Assays (SPA) by Amersham. Alternatively, the interaction of the proteins of the invention with cellular proteins could be the basis for a cell-based screening assay, in which both proteins are fluorescently labeled and interaction of both proteins is detected by analysing cotranslocation of both proteins with a cellular imaging reader, as has been developed for example, but not exclusively, by Cellomics or EvotecOAI. In all cases the two or more binding partners can be different proteins with one being the protein of the invention, or in case of dimerization and/or oligomerization the protein of the invention itself. Proteins of the invention, for which one target mechanism of interest, but not the only one, would be such protein/protein interactions are PSAP, SFTPB, and/or FLJ40379.

Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups.

The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups.

5 Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for
10 random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and
15 compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening
20 assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal.

Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to
25 the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as plastic pins or some other surface. The test
30 compounds are reacted with the proteins or fragments thereof, and washed. Bound proteins are then detected by methods well known in the art. Purified proteins can also be coated directly onto plates for use in the

aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support. In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding the protein
5 specifically compete with a test compound for binding the protein. In this manner, the antibodies can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the protein.

The nucleic acids encoding the proteins of the invention can be used to
10 generate transgenic cell lines and animals. These transgenic non-human animals are useful in the study of the function and regulation of the proteins of the invention in vivo. Transgenic animals, particularly mammalian transgenic animals, can serve as a model system for the investigation of many developmental and cellular processes common to
15 humans. A variety of non-human models of metabolic disorders can be used to test modulators of the protein of the invention. Misexpression (for example, overexpression or lack of expression) of the protein of the invention, particular feeding conditions, and/or administration of biologically active compounds can create models of metabolic disorders.

20

In one embodiment of the invention, such assays use mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor) mice). Such mice develop typical symptoms of diabetes, show hepatic lipid
25 accumulation and frequently have increased plasma lipid levels (see Bruning et al, 1998, Mol. Cell. 2:449-569). Susceptible wild-type mice (for example C57Bl/6) show similar symptoms if fed a high fat diet. In addition to testing the expression of the proteins of the invention in such mouse strains (see EXAMPLE 4), these mice could be used to test whether
30 administration of a candidate modulator alters for example lipid accumulation in the liver, in plasma, or adipose tissues using standard

assays well known in the art, such as FPLC, colorimetric assays, blood glucose level tests, insulin tolerance tests and others.

Transgenic animals may be made through homologous recombination in embryonic stem cells, where the normal locus of the gene encoding the protein of the invention is mutated. Alternatively, a nucleic acid construct encoding the protein is injected into oocytes and is randomly integrated into the genome. One may also express the genes of the invention or variants thereof in tissues where they are not normally expressed or at abnormal times of development. Furthermore, variants of the genes of the invention like specific constructs expressing anti-sense molecules or expression of dominant negative mutations, which will block or alter the expression of the proteins of the invention may be randomly integrated into the genome. A detectable marker, such as lac Z or luciferase may be introduced into the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detectable change in phenotype. Vectors for stable integration include plasmids, retroviruses and other animal viruses, yeast artificial chromosomes (YACs), and the like. DNA constructs for homologous recombination will contain at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. Conveniently, markers for positive and negative selection are included. DNA constructs for random integration do not need to contain regions of homology to mediate recombination. DNA constructs for random integration will consist of the nucleic acids encoding the proteins of the invention, a regulatory element (promoter), an intron and a poly-adenylation signal. Methods for generating cells having targeted gene modifications through homologous recombination are known in the field. For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer and are grown in the presence of leukemia inhibiting factor (LIF). ES

or embryonic cells may be transfected and can then be used to produce transgenic animals. After transfection, the ES cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be selected by employing a selection medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination. Colonies that are positive may then be used for embryo manipulation and morula aggregation. Briefly, morulae are obtained from 4 to 6 week old superovulated females, the Zona Pellucida is removed and the morulae are put into small depressions of a tissue culture dish. The ES cells are trypsinized, and the modified cells are placed into the depression closely to the morulae. On the following day the aggregates are transferred into the uterine horns of pseudopregnant females. Females are then allowed to go to term. Chimeric offsprings can be readily detected by a change in coat color and are subsequently screened for the transmission of the mutation into the next generation (F1-generation). Offspring of the F1-generation are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc., for example, mouse, rat, guinea pig, sheep, cow, pig, and others. The transgenic animals may be used in functional studies, drug screening, and other applications and are useful in the study of the function and regulation of the proteins of the invention in vivo.

Finally, the invention also relates to a kit comprising at least one of

- (a) a nucleic acid molecule coding for a protein of the invention or a fragment thereof;
- (b) a protein of the invention or a fragment or an isoform thereof;
- (c) a vector comprising the nucleic acid of (a);
- (d) a host cell comprising the nucleic acid of (a) or the vector of (b);

- (e) a polypeptide encoded by the nucleic acid of (a);
- (f) a fusion polypeptide encoded by the nucleic acid of (a);
- (g) an antibody, an aptamer or another effector against the nucleic acid of (a) or the polypeptide of (b), (e) or (f) and
- 5 (h) an anti-sense oligonucleotide of the nucleic acid of (a).

The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

10

The Figures show:

Figure 1 shows the content of energy storage metabolites (ESM; triglyceride (TG) and glycogen) of *Drosophila* Saposin-related (GadFly Accession Number CG12070) mutants. Shown is the change of triglyceride content of HD-EP(3)36824 flies caused by integration of the P-vector into the annotated transcription unit ('HD-36824 (TG, 70°C)', column 3 and 'HD-36824 (TG, 90°C)', column 6) in comparison to controls containing about 880 fly lines of the proprietary EP collection ('HD-control (TG, 70°C)'), column 1) and wild-type controls determined in 4 independent assays (referred to as 'WT-control (TG, 70°C)', column 2), and in comparison to controls containing about 2100 fly lines of the proprietary EP collection ('HD-control (TG, 90°C)'), column 4) and wild-type controls determined in more than 80 independent assays (referred to as 'WT-control (TG, 90°C)', column 5). Also shown is the change of glycogen content of HD-EP(3)36824 flies caused by integration of the P-vector the into the annotated transcription unit ('HD-36824 (glycogen, 90°C)', column 8) in comparison to an internal assay control including two wild-type strains and a one HD-line (referred to as 'assay control (glycogen, 90°C)' column 7).

Figure 2 shows the molecular organization of the mutated Saposin-related (Sap-r; GadFly Accession Number CG12070) gene locus.

Figure 3 shows human saposin-related proteins: Prosaposin (PSAP), human surfactant, pulmonary-associated protein B (SFTPB), or human hypothetical protein FLJ40379

Figure 3A shows the nucleic acid sequence of human PSAP (SEQ ID NO: 1)

Figure 3B shows the amino acid sequence (one-letter code) of human PSAP (SEQ ID NO: 2)

Figure 3C shows the nucleic acid sequence of human SFTPB (SEQ ID NO: 3)

Figure 3D shows the amino acid sequence (one-letter code) of human SFTPB (SEQ ID NO: 4)

Figure 3E shows the nucleic acid sequence of human FLJ40379 (SEQ ID NO: 5)

Figure 3F shows the amino acid sequence (one-letter code) of human FLJ40379 (SEQ ID NO: 6)

Figure 3G shows the comparison (ClustalW (1.83) protein sequence alignment analysis) of human and Drosophila protein. Gaps in the alignment are represented as -. In the figure 'PSAP Hs' refers to human prosaposin, 'FLJ40379 Hs' refers to human hypothetical protein FLJ40379, 'SFTPB Hs' refers to human surfactant, pulmonary-associated protein B, and 'Sap-r Dm' refers to Drosophila Saposin-related protein .

Figure 4 shows the quantitative analysis of prosaposin (Psap) expression in mammalian tissues. The relative RNA-expression is shown on the Y-axis, in Figure 4A to 4C the tissues tested are given on the X-axis. WAT refers to white adipose tissue, BAT refers to brown adipose tissue. In Figure 4D, the X-axis represents the time axis. 'd0' refers to day 0 (start of the experiment), 'd2' - 'd10' refers to day 2 - day 10 of adipocyte differentiation).

Figure 4A shows the quantitative analysis of Psap expression in mouse wild-type tissues.

Figure 4B shows the quantitative analysis of Psap expression in wild-type mice (wt-mice), compared to genetically obese mice (ob/ob-mice) and to fasted mice (fasted-mice).

Figure 4C shows the quantitative analysis of Psap expression in mice fed with a control diet compared to mice fed with a high fat diet.

Figure 4D shows the quantitative analysis of Psap expression in mammalian fibroblast (3T3-L1) cells, during the differentiation from preadipocytes to mature adipocytes.

Figure 5 shows the quantitative analysis of prosaposin expression in human abdominal adipocyte cells, during the differentiation from preadipocytes to mature adipocytes.

The examples illustrate the invention:

Example 1: Measurement of energy storage metabolites (ESM) contents in *Drosophila*

Mutant flies are obtained from a fly mutation stock collection. The flies are grown under standard conditions known to those skilled in the art. In the course of the experiment, additional feedings with bakers yeast (*Saccharomyces cerevisiae*) are provided for the EP-line HD-EP(3)36824. The average change of triglyceride and glycogen (herein referred to as energy storage metabolites, ESM) content of *Drosophila* containing the EP-vectors as homozygous viable integration was investigated in comparison to control flies, respectively (see Figure 1). For determination of ESM content, flies were incubated for 5 min at 70°C or 90°C in an aqueous buffer using a waterbath, followed by hot extraction. After another 5 min incubation at 70°C or 90°C and mild centrifugation, the

triglyceride content of the flies extract was determined using Sigma Triglyceride (INT 336-10 or -20) assay by measuring changes in the optical density according to the manufacturer's protocol, and the glycogen content of the flies extract was determined using Roche (Starch UV-method Cat. No. 0207748) assay by measuring changes in the optical density according to the manufacturer's protocol. As a reference the protein content of the same extract was measured using BIO-RAD DC Protein Assay according to the manufacturer's protocol. These experiments and assays were repeated several times.

The average triglyceride level of 883 fly lines of the proprietary EP-collection determined at 70°C (referred to as 'HD-control (TG, 70°C)') is shown as 100% in the first column in Figure 1. The average triglyceride level of *Drosophila* wild-type strain Oregon R flies determined in 4 independent assays at 70°C (referred to as 'WT-control (TG, 70°C)') is shown as 116% in the second column in Figure 1. The average triglyceride level (μg triglyceride/ μg protein) of 2108 fly lines of the proprietary EP-collection determined at 90°C (referred to as 'HD-control (TG, 90°C)') is shown as 100% in the fourth column in Figure 1. The average triglyceride level (μg triglyceride/ μg protein) of *Drosophila* wild-type strain Oregon R flies determined in 84 independent assays at 90°C (referred to as 'WT-control (TG, 90°C)') is shown as 102% in the fifth column of Figure 1. The average glycogen level (μg glycogen/ μg protein) of an internal assay control consisting of two different wild-type strains and an inconspicuous EP-line of the HD stock collection (referred to as 'assay control (glycogen, 90°C)') is shown as 100% in the seventh column in Figure 1. The average triglyceride level (μg triglyceride/ μg protein) of all flies of the EP collection (referred to as 'EP-control') is shown as 100% in the first column in Figure 5. Standard deviations of the measurements are shown as thin bars.

HD-EP(3)36824 homozygous flies show constantly a lower triglyceride content than the controls (column 3 in Figure 1, 'HD-36824 (TG, 70°C)'; column 6 in Figure 1, 'HD-36824 (TG, 90°C)'). HD-EP(2)21554 homozygous flies also show a lower glycogen content than the controls (column 8 in Figure 1, 'HD-36824 (glycogen, 90°C)'). Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage metabolites.

Example 2: Identification of *Drosophila* genes responsible for changes in metabolite contents

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(3)36824) integration. Using the isolated genomic sequences from HD-EP(3)36824 homozygous flies public databases like Berkeley *Drosophila* Genome Project (GadFly) were screened. The chromosomal localization site of HD-EP(3)36824 vector integration is at gene locus 3R, 100A6-7 (Flybase and Gadfly). The homozygous viable integration site of the HD-EP(3)36824 vector into base pair 535 of the Sap-r transcript CG12070-RA and 61 base pairs 5prime of the Sap-r transcript CG12070-RB in antisense orientation was confirmed. Therefore, expression of the cDNAs encoded by Sap-r could be affected by integration of the vector of line HD-EP(3)36824 leading to a change in the amount of energy storage metabolites. Figure 2 shows the molecular organization of this gene locus. A black double arrow in middle of the Figure represents the genomic DNA sequence. The space between two ticks represents a stretch of 1000 base pairs. The black triangle labeled 'HD-EP36824' indicates the integration site of the EP-vector. The transcripts of Sap-r (as predicted by the Berkeley *Drosophila* Genome Project) are shown as dark gray bars (exons) linked by dark gray lines (introns) in the lower half of the Figure labeled as 'Sap-r'.

Table 1 is summarizing the data of our molecular analysis of the *Drosophila* proteins identified in this invention as being involved in the regulation of the metabolism.

5

Table 1. Molecular analysis of *Drosophila* Saposin-related

	Analysis	Result
10	Protein domains	Beta-Ig-H3/Fasciclin (Flybase)
	InterPro analysis	Saposin type B, Saposin type A, Surfactant protein B
	Locus	3R, 100A6-7 (Flybase, Gd-fly release 3)
	cDNA (sap-r)	AI108030 (578 base pairs mRNA), AI109190 (635 base pairs mRNA)
	genomic DNA	AE003775
15	RefSeq (Sap-r)	NM_079858, NM_170529, NP_524597, NP_733408
	<i>Drosophila</i> mutations & mutants	Homozygotes for deficiencies removing Sap-r are fertile and viable and show no obvious phenotype. (Flybase)

20

Example 3: Identification of the human Saposin-related homologous proteins

Saposin-related homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or
25 birds, preferably from humans, mouse, or *Drosophila*. Sequences homologous to *Drosophila* Saposin-related were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI) (see,
30 Altschul et al., 1997, Nucleic Acids Res. 25:3389-3402). Table 2 shows preferred human homologs of the *Drosophila* Saposin-related gene.

The term "polynucleotide comprising the nucleotide sequence as shown in GenBank Accession number" relates to the expressible gene of the nucleotide sequences deposited under the corresponding GenBank Accession number. The term "GenBank Accession number" relates to NCBI GenBank database entries (Ref.: Benson et al., Nucleic Acids Res. 28 (2000) 15-18).

Table 2. Human proteins homologous to Drosophila Saposin-related proteins

I. Human proteins homologous to Drosophila Sap-r

IA. Human prosaposin (PSAP)

NCBI human locus identification (ID): 5660, Hs PSAP, prosaposin (variant Gaucher disease and variant metachromatic leukodystrophy), 10q21-q22
RefSeq: GenBank Accession Number: NM_002778

IB. Human surfactant, pulmonary-associated protein B (SFTPB)

NCBI human locus identification (ID): 6439; Hs SFTPB, surfactant, pulmonary-associated protein B, 2p12-p11.2

Aliases: SP-B, PSP-B, SFTB3, SFTP3

RefSeq: GenBank Accession Number: NM_000542

IC. Human hypothetical protein FLJ40379

Nucleotide: ENSEMBL Accession Number: ENSG00000173005

Protein: ENSEMBL Accession Number: ENSP00000308224, International Protein Index (IPI) Accession Number: IPI00163920

Example 4: Identification of secreted factors expressed in pancreas

A screen for secreted factors expressed in developing mouse pancreas was carried out according to methods known by those skilled in the art (see, for example Pera E.M. and De Robertis E.M., (2000) Mech Dev 96(2): 183-195) with several modifications.

Expression cDNA library:

A mouse embryonic stage 9.5-15 pancreatic bud library was prepared in pCMVSPORT-6 vector using SUPERScript Plasmid System from Invitrogen (cat.#18248) according to the manufacturer's instructions. The non-amplified library was electroporated into MaxEff DH10B cells (Invitrogen).

Secretion cloning

Bacterial clones from agar plates were picked with sterile toothpicks and cultured in 96-deep-well microtiter plates in 1 ml of LB-ampicillin (see Sambrook et al., supra). Aliquots of 8 cultures were pooled, and plasmid DNAs isolated by BioRobot_9600 (QIAGEN) using QIAprep(r) Turbo BioRobot Kit (QIAGEN). Human 293 cells were cultured in 75 ml tissue culture flasks in DMEM and 10% fetal calf serum. At 90-99% confluence, the cells were splitted at 1:3 ratio and plated at poly-D-lysine (Sigma) coated 96-well plates, 150 μ l/well. Next day the the cells were transfected with 100-500 ng plasmids using lipofectamine 2000 (Invitrogen), 0,5 μ l/well, in 100 μ l/well Optimem (Invitrogen). After 6 h the medium was exchanged for fresh complete growth medium. 24 hours after transfection, the cells were washed twice with DMEM w/o Cysteine and Methionine (Invitrogen), supplemented with 1% dialysed Bovine serum (Sigma) with 50 mg/ml Heparin (Sigma) and glutamine. The cells were labeled in 50 μ l/well of the same medium and 0,75 μ l/well (-S35 Met-label (HARTMANN ANALYTIC GmbH, #44138). 12 hours later, 10 μ l aliquots of the supernatants were harvested in 96-well PCR plates and subjected to SDS

gel electrophoresis in precast $4 \pm 20\%$ gradient polyacrylamide Criterion gels (Biorad) under reducing conditions, using Criterion Dodeca Cell gel running chamber (Biorad). The gels were fixed in 10% acetic acid, 25% isopropanol for 30 min, soaked 15-30 min in AMPLIFY reagent (Amersham), dried and exposed to X-OMAT (AR) film (Kodak). Positive clones were identified by sub-selection. The 8 individual bacterial clones of each positive pool were regrown in 96-well-plates, DNA of individual clones was prepared and used for transfection as described. If one of the clones yielded proteins of the same size as that of the original pool, a positive clone was identified. Positive clones were partially sequenced from the 5' end (SEQLAB, Goettingen). Sequences (of about 500 nucleotides) were compared to public nucleotide and protein databases to reveal similarity to previously described proteins.

Example 5: Expression of the polypeptides in mammalian tissues

To analyse the expression of the polypeptides disclosed in this invention in mammalian tissues, several mouse strains (preferably mice strains C57Bl/6J, C57Bl/6 ob/ob and C57Bl/KS db/db which are standard model systems in obesity and diabetes research) were purchased from Harlan Winkelmann (33178 Borcheln, Germany) and maintained under constant temperature (preferably 22°C), 40 per cent humidity and a light / dark cycle of preferably 14 / 10 hours. The mice were fed a standard chow (for example, from ssniff Spezialitäten GmbH, order number ssniff M-Z V1126-000). For the fasting experiment ("fasted wild-type mice"), wild-type mice were starved for 48 h without food, but only water supplied ad libitum (see, for example, Schnetzler et al. J Clin Invest 1993 Jul;92(1):272-80, Mizuno et al. Proc Natl Acad Sci U S A 1996 Apr 16;93(8):3434-8). Animals were sacrificed at an age of 6 to 8 weeks. The animal tissues were isolated according to standard procedures known to

those skilled in the art, snap frozen in liquid nitrogen and stored at -80°C until needed.

For analyzing the role of the proteins disclosed in this invention in the in vitro differentiation of different mammalian cell culture cells for the conversion of pre-adipocytes to adipocytes, mammalian fibroblast (3T3-L1) cells (e.g., Green & Kehinde, Cell 1: 113-116, 1974) were obtained from the American Tissue Culture Collection (ATCC, Hanassas, VA, USA; ATCC- CL 173). 3T3-L1 cells were maintained as fibroblasts and differentiated into adipocytes as described in the prior art (e.g., Qiu. et al., J. Biol. Chem. 276:11988-95, 2001; Sliker et al., BBRC 251: 225-9, 1998). In brief, cells were plated in DMEM/10% FCS (Invitrogen, Karlsruhe, Germany) at 50,000 cells/well in duplicates in 6-well plastic dishes and cultured in a humidified atmosphere of 5% CO₂ at 37°C. At confluence (defined as day 0: d0) cells were transferred to serum-free (SF) medium, containing DMEM/HamF12 (3:1; Invitrogen), fetuin (300 µg/ml; Sigma, Munich, Germany), transferrin (2 µg/ml; Sigma), pantothenate (17 µM; Sigma), biotin (1 µM; Sigma), and EGF (0.8 nM; Hoffmann-La Roche, Basel, Switzerland). Differentiation was induced by adding dexamethasone (DEX; 1 µM; Sigma), 3-methyl-isobutyl-1-methylxanthine (MIX; 0.5 mM; Sigma), and bovine insulin (5 µg/ml; Invitrogen). Four days after confluence (d4), cells were kept in SF medium, containing bovine insulin (5 µg/ml) until differentiation was completed. At various time points of the differentiation procedure, beginning with day 0 (day of confluence) and day 2 (hormone addition; for example, dexamethasone and 3-isobutyl-1-methylxanthine), up to 10 days of differentiation, suitable aliquots of cells were taken every two days.

RNA was isolated from mouse tissues or cell culture cells using Trizol Reagent (for example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in combination with an DNase-treatment according to the instructions of the

manufacturers and as known to those skilled in the art. Total RNA was reverse transcribed (preferably using Superscript II RNaseH- Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the Taqman 2xPCR Master Mix (from Applied Biosystems, Weiterstadt, Germany; the Mix contains according to the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence Detection System (from Applied Biosystems, Weiterstadt, Germany).

Example 6. Analysis of the differential expression of transcripts of the proteins of the invention in human tissues

RNA preparation from human primary adipose tissues was done as described in EXAMPLE 2. The hybridization and scanning was performed as described in the manufactures manual (see Affymetrix Technical Manual, 2002, obtained from Affmetrix, Santa Clara, USA).

The expression analysis (using Affymetrix GeneChips) of the psap gene using primary human abdominal adipocyte differentiation clearly shows differential expression of human psap in adipocytes. Two independent experiments were done. Both experiments show that the psap transcripts are the most abundance at day 0 compared to day 12 during differentiation. These data further confirm the mouse 3T3L-1 differentiation data.

Thus, the psap protein has to be significantly decreased in order for the preadipocytes to differentiate into mature adipocyte. The psap protein in preadipocytes has the potential to inhibit adipose differentiation. Therefore, the psap protein might play an essential role in the regulation of

human metabolism, in particular in the regulation of adipogenesis and thus it might be an essential role in obesity, diabetes, and/or metabolic syndrome.

Claims

1. A pharmaceutical composition comprising a nucleic acid molecule encoding Saposin-related or a homologue thereof or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic acid molecule or said polypeptide, preferably together with pharmaceutically acceptable carriers and diluents.
2. The composition of claim 1, wherein the nucleic acid molecule is a vertebrate or insect Saposin-related nucleic acid, particularly encoding the human homologs (such as human prosaposin (PSAP), human pulmonary surfactant-associated protein B (SFTPB), or human hypothetical protein FLJ40379), and/or a nucleic molecule which is complementary thereto or a fragment thereof or a variant thereof.
3. The composition of claim 1 or 2, wherein said nucleic acid molecule is selected from the group consisting of
- (a) a nucleic acid molecule encoding a polypeptide as shown in SEQ ID NO: 2, 4, and/or 6, or an isoform, fragment or variant of the polypeptide as shown in SEQ ID NO: 2, 4, and/or 6 and/or a nucleic acid molecule complementary thereto;
 - (b) a nucleic acid molecule which comprises or is the nucleic acid molecule as shown in SEQ ID NO: 1, 3, and/or 5 and/or a nucleic acid molecule complementary thereto;
 - (c) a nucleic acid molecule being degenerate with as a result of the genetic code to the nucleic acid sequences as defined in (a) or (b),
 - (d) a nucleic acid molecule that hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a nucleic acid molecule

as defined in claim 2 or as defined in (a) to (c) and/or a nucleic acid molecule which is complementary thereto;

(e) a nucleic acid molecule that encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the human PSAP, SFTPB, and/or FLJ40379, as defined in claim 2 or to a polypeptide as defined in (a);

(f) a nucleic acid molecule that differs from the nucleic acid molecule of (a) to (e) by mutation and wherein said mutation causes an alteration, deletion, duplication or premature stop in the encoded polypeptide.

4. The composition of any one of claims 1-3, wherein the nucleic acid molecule is a DNA molecule, particularly a cDNA or a genomic DNA.

5. The composition of any one of claims 1-4, wherein said nucleic acid encodes a polypeptide contributing to regulating the energy homeostasis and/or the metabolism of triglycerides.

6. The composition of any one of claims 1-5, wherein said nucleic acid molecule is a recombinant nucleic acid molecule.

7. The composition of any one of claims 1-6, wherein the nucleic acid molecule is a vector, particularly an expression vector.

8. The composition of any one of claims 1-5, wherein the polypeptide is a recombinant polypeptide.

9. The composition of claim 8, wherein said recombinant polypeptide is a fusion polypeptide.

10. The composition of any one of claims 1-7, wherein said nucleic acid molecule is selected from hybridization probes, primers and anti-sense oligonucleotides.
- 5 11. The composition of any one of claims 1-10 which is a diagnostic composition.
12. The composition of any one of claims 1-10 which is a therapeutic composition.
- 10 13. The composition of any one of claims 1-12 for the manufacture of an agent for detecting and/or verifying, for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome, obesity, diabetes
15 mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and/or gallstones, and others, in cells, cell masses, organs and/or subjects.
- 20 14. Use of a nucleic acid molecule of the PSAP, SFTPB, and/or FLJ40379 gene family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic or polypeptide for controlling the function of a gene and/or a gene product which is influenced and/or modified by a Saposin-related homologous
25 polypeptide.
- 30 15. Use of the nucleic acid molecule of the PSAP, SFTPB, and/or FLJ40379 gene family or use of a nucleic acid molecule encoding Saposin-related or a homologue thereof or use of a polypeptide encoded thereby, or use of a fragment or a variant of said nucleic acid molecule or said polypeptide, or use of an effector of said nucleic acid molecule or said polypeptide for identifying substances

capable of interacting with a Saposin-related homologous polypeptide.

5 16. A non-human transgenic animal exhibiting a modified expression of a Saposin-related homologous polypeptide.

17. The animal of claim 16, wherein the expression of the Saposin-related homologous polypeptide is increased and/or reduced.

10 18. A recombinant host cell exhibiting a modified expression of a Saposin-related homologous polypeptide, or a recombinant host cell which comprises a nucleic acid molecule as defined in any one of claims 1 to 6.

15 19. The cell of claim 18 which is a human cell.

20 20. A method of identifying a (poly)peptide involved in the regulation of energy homeostasis and/or metabolism of triglycerides in a mammal comprising the steps of

- 25 (a) contacting a collection of (poly)peptides with a Saposin-related homologous polypeptide or a fragment thereof under conditions that allow binding of said (poly)peptides;
- (b) removing (poly)peptides which do not bind and
- (c) identifying (poly)peptides that bind to said Saposin-related homologous polypeptide.

30 21. A method of screening for an agent which modulates the interaction of a Saposin-related homologous polypeptide with a binding target/agent, comprising the steps of

- (a) incubating a mixture comprising

(aa) a Saposin-related homologous polypeptide or a fragment thereof;

(ab) a binding target/agent of said Saposin-related homologous polypeptide or fragment thereof; and

(ac) a candidate agent

under conditions whereby said Saposin-related homologous polypeptide or fragment thereof specifically binds to said binding target/agent at a reference affinity;

(b) detecting the binding affinity of said Saposin-related homologous polypeptide or fragment thereof to said binding target to determine a (candidate) agent-biased affinity; and

(c) determining a difference between (candidate) agent-biased affinity and reference affinity.

22. A method for screening for an agent, which modulates the activity of a Saposin-related homologous polypeptide, comprising the steps of

(a) incubating a mixture comprising

(aa) a Saposin-related homologous polypeptide or a fragment thereof; and

(ab) a candidate agent

under conditions whereby said Saposin-related homologous polypeptide or fragment thereof exhibits a reference activity,

(b) detecting the activity of said Saposin-related homologous polypeptide or fragment thereof to determine a (candidate) agent-biased activity; and

(c) determining a difference between (candidate) agent-biased activity and reference activity.

23. A method of producing a composition comprising the (poly)peptide identified by the method of claim 20 or the agent identified by the

method of claim 21 or 22 with a pharmaceutically acceptable carrier, diluent and/or adjuvant.

- 5 24. The method of claim 23 wherein said composition is a pharmaceutical composition for preventing, alleviating or treating of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, such as metabolic syndrome, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hyper-cholesterolemia (dyslipidemia), and / or gallstones, and other diseases and disorders.
- 10
25. Use of a (poly)peptide as identified by the method of claim 20 or of an agent as identified by the method of claim 21 or 22 for the preparation of a pharmaceutical composition for the treatment, alleviation and/or prevention of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hyper-cholesterolemia (dyslipidemia), and / or gallstones, and other diseases and disorders.
- 15
26. Use of a nucleic acid molecule as defined in any one of claims 1 to 6 or 10, use of a polypeptide as defined in any one of claims 1 to 6, 8 or 9, use of a vector as defined in claim 7, use of a host cell as defined in claim 18 or 19 for the preparation of a pharmaceutical composition for the treatment, alleviation and/or prevention of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and / or gallstones, and other diseases and disorders.
- 20
- 25
- 30

27. Use of a nucleic acid molecule of the PSAP, SFTPb, and/or FLJ40379 gene family or of a fragment thereof for the preparation of a non-human animal which over- or under-expresses the PSAP, SFTPb, and/or FLJ40379 gene product.
28. Kit comprising at least one of
- (a) a PSAP, SFTPb, and/or FLJ40379 nucleic acid molecule or a fragment or an isoform thereof;
 - (b) a PSAP, SFTPb, and/or FLJ40379 amino acid molecule or a fragment or an isoform thereof;
 - (c) a vector comprising the nucleic acid of (a);
 - (d) a host cell comprising the nucleic acid of (a) or the vector of (b);
 - (e) a polypeptide encoded by the nucleic acid of (a), expressed by the vector of (c) or the host cell of (a);
 - (f) a fusion polypeptide encoded by the nucleic acid of (a);
 - (g) an antibody, an aptamer or another effector against the nucleic acid of (a) or the polypeptide of (b), (e), or (f) and /or
 - (h) an anti-sense oligonucleotide of the nucleic acid of (a).

Abstract

5 The present invention discloses Saposin-related homologous proteins
regulating the energy homeostasis and the metabolism of triglycerides, and
polynucleotides, which identify and encode the proteins disclosed in this
invention. The invention also relates to the use of these sequences in the
diagnosis, study, prevention, and treatment of metabolic diseases and
disorders.

10

ld/kt 26.03.2003

26 März 2003

EPO - Munich
73

26. März 2003

Figure 1. Energy Storage Metabolites (ESM) content of a *Drosophila* Saposin-related (Sap-r, GdFly Accession Number CG12070) mutant

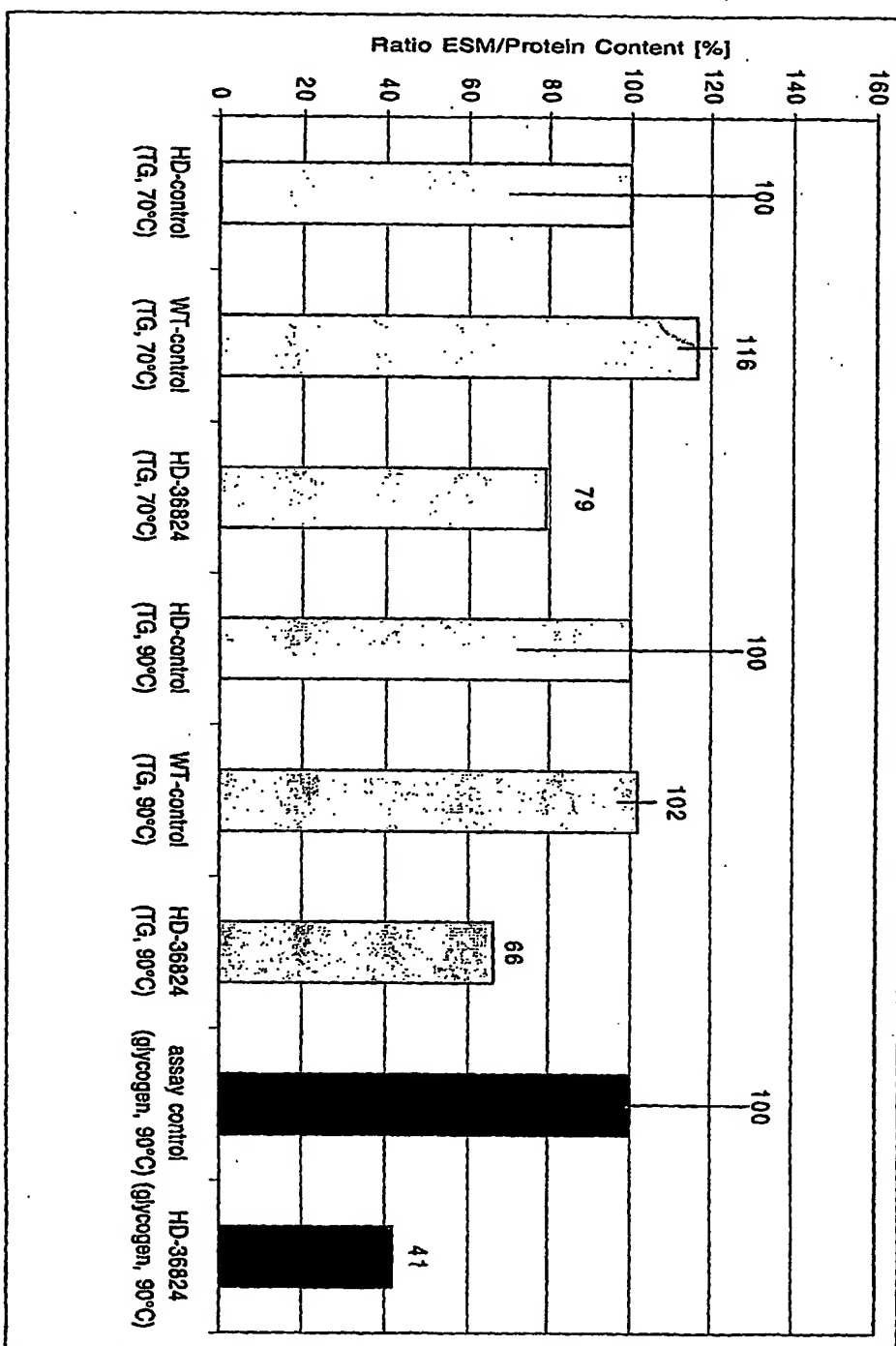


Figure 2. Molecular organization of the *Drosophila* Saposin-related gene (GadFly Accession Number CG12070)

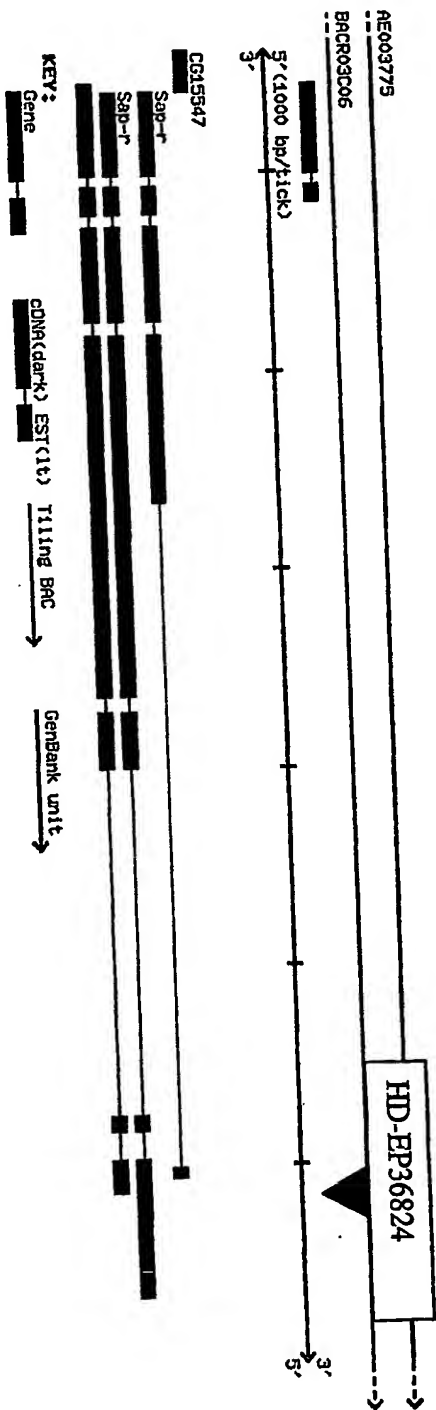


Figure 3. Human Prosaposin (PSAP); surfactant, human pulmonary-associated protein B (SFTPB); and human hypothetical protein FLJ 40379

Figure 3A. Human prosaposin (variant Gaucher disease and variant metachromatic leukodystrophy, PSAP), Nucleic acid sequence (SEQ ID NO:1)

```

1  gggcgggcgc attgcagact ggggagtcag acgggtgctat gtacgcccctc ttcctcctgg
61  ccagcctcct gggcgcggct ctagccggcc cggtccttgg actgaaagaa tgcaccaggg
121 gctcggcagt gtggtgccag aatgtgaaga cggcgctccga ctgcggggca gtgaagcact
181 gcctgcagac cgtttggaac aagccaacag tgaaatccct tccctgcgac atatgcaaag
241 acgttgctcac cgcagctggt gatatgctga aggacaatgc cactgaggag gagatccttg
301 tttacttgga ctcctacctc cctgtcatcc cgaaacggaa catgtctgct tcatgcaagg
361 agatagtggg gtgctctgct ctcaacctct tggacatcat taaaggagaa atgagccgtc
421 ctggggaggt gtgctctgct ctcaacctct gcgagtctct ccagaagcac ctagcagagc
481 tgaatcacca gaagcagctg gagtccaata agatcccaga gctggacatg actgaggttg
541 tggccccctt catggccaac atccctctcc tcccttacct tcaggacggc ccccgagca
601 agccccagcc aaaggataat ggggacggtt gccaggactg cattcagatg gtgactgaca
661 tccagactgc tgtacggacc aactccacct ttgtccaggc cttggtggaa catgtcaagg
721 aggagtgtga ccgcctgggc cctggcatgg ccgacatatt caagaactat atcagccagt
781 attctgaaat tgctatccag atgatgatgc acatgcaacc caaggagatc tgtgcgctgg
841 ttgggttctg tgatgaggtg aaagatgatg ccatgcagac tctggtcccc gccaaagtgg
901 cctccaagaa tgtcatccct gccctggaac tgggtggagcc cattaaagaag cagcaggtcc
961 cagcaaagtc tgatgtttac tgtgaggtgt gtgaattcct ggtgaaggag gtgaccaagc
1021 tgattgacaa caacaagact gagaaagaaa tactcgacgc ttttgacaaa atgtgctcga
1081 agctgccgaa gtccctgtcg gaagagtgcc aggaggtggt ggacacgtac ggcagctcca
1141 tcctgtccat cctgctggag gaggtcagcc ctgagctggt gtgcagcatg ctgcacctct
1201 gctctggcac gcggtgcct gcaactgacc ttacgtgac tcagccaaag gacggtggct
1261 tctgcaagt gtgcaagaag ctggtgggtt atttggatcg caacctggag aaaaacagca
1321 ccaagcagga gatcctggct gctcttgaga aaggctgcag ctctctgcca gaccttacc
1381 agaagcagtg tgatcagttt gtggcagagt acgagcccgt gctgatcgag atcctggtgg
1441 aggtgatgga tccttccttc gtgtgcttga aaattggagc ctgccccctc gcccataagc
1501 ccttggttgg aactgagaag tgtatatggg gcccaagcta ctggtgccag aacacagaga
1561 cagcagccca gtgcaatgct gtcagacatt gcaaacgcca tgtgtggaac taggaggagg
1621 aatattccat cttggcagaa accacagcat tgggtttttt ctacttgtgt gtctggggga
1681 atgaacgcac agatctgttt gactttgtta taaaaatagg gctccccac ctccccatt
1741 tctgtgtcct ttattgtagc attgctgtct gcaagggagc ccctagcccc tagccccctg
1801 cagacatagc tgcttcagtg ccccttttct ctctgctaga tggatgttga tgcactggag
1861 gtcttttagc ctgcccttgc atggcgccct ctggaggagg agagagctct gctggcatga
1921 gccacagttt cttgactgga ggccatcaac cctcttggtt gaggccttgt tctgagccct
1981 gacatgtgct tgggcactgg tgggcctggg cttctgaggt ggccctctgc cctgatcagg
2041 gacctcccc gctttccttg gccctcagat tgaacaaagc agcaaaacaa aggcagtttt
2101 atatgaaaga ttagaagcct ggaataatca ggctttttta atgatgtaat tcccactgta
2161 atagcatagg gattttggaa gcagctgctg gtggccttgg acatcagttg ggccaagggt
2221 tctctgtccc tggttcaact gtgatttggc tttcccgtgt ctttccctgg gatgccttgt
2281 ttgggggttct gtgggttttg gtgggaagag ggccatctgc ctgaatgtaa cctgctagct
2341 ctccgaagcc ctgcgggcct ggcttgtgtg agcgtgtgga cagtgggtgg cgcgctgtgc
2401 ctgctcgtgt tgcctacatg tccctggctt ttgagggcgt gcttcagcct gcacctctcc
2461 cttgtctcat agatgtcctt tttgaccttt tcaaataaat atggatggcg agctcctagg
2521 cctctggctt cctggtagag ggccgcatgc cgaagggtct gtctgggtgt ggtattgtag
2581 ctgggggtgt ggggggttga agctgtctgt ggccacttgg ggcaccacg cttctgtcca
2641 cttctggttg ccaggagaca gcaagcaaa ccagcaggac atgaagttgc tattaatagg
2701 acttcgtgat tttgttttg cactaaagtt tctgtgattt aacaataaaa tctctgttag
2761 cccccgc

```

Figure 3B. Human prosaposin (sphingolipid activator protein-1), amino acid sequence (SEQ ID NO: 2).

```

1 myalfillasl lgaalagpvl glkectrgsa vwcqnvktas dcgavkhclq tvwnkptvks
61 lpcdickdvv taagdmlkdn ateeeilvyl ektcdwlpkp nmsasckeiv dsylpvildi
121 ikgemsrpge vcsalnices lqkhlaelnh qkqlesnkip eldmtevvap fmanipilly
181 pqdgprskpq pkdngdvcqd ciqmvtdiqt avrtntstfvq alvehvkeec drlgpgmadi
241 cknyisqyse iaigmmmmhmq pkeicalvgf cdevkempmq tlpakvask nvipalelve
301 pikkhevpak sdvycevef lvkevtklid nnktekeild afdkmcsklp kslseecgev
361 vdytgssils illeevspel vcsmlhlcsq trlpaltvhv tqpkdggfce vckklvgylid
421 rnlekntskq eilaalekgc sflpdpyqkq cdqfvaeyep vlieilvevm dpsfvclkg
481 acpsahkpll gtekciwgps ywcqntetaa qcnavehckr hvwn

```

Figure 3C. Human surfactant, pulmonary-associated protein B (SFTPB), nucleic acid sequence (SEQ ID NO: 3)

```

1 atggctgagt cacacctgct gcagtggttg ctgctgctgc tgcccacgct ctgtggccca
61 ggcactgctg cctggaccac ctcctccttg gcctgtgccc agggccctga gttctggtgc
121 caaagcctgg agcaagcatt gcagtcgaga gccctagggc attgcctaca ggaagtctgg
181 ggacatgtgg gagccgatga cctatgcaca gagtgtgagg acatcgctca catccttaac
241 aagatggcca aggaggccat ttccaggac acgatgagga agttcctgga gcaggagtgc
301 aacgtcctcc ccttgaagct gctcatgccc cagtgcaccc aagtgcctga cgactacttc
361 cccctggtca tcgactactt ccagaaccag attgactcaa acggcatctg tatgcacctg
421 ggcctgtgca aatcccggca gccagagcca gagcaggagc cagggatgtc agaccccctg
481 cccaaacctc tgccgggacc tctgccagac cctctgctgg acaagctcgt cctccctgtg
541 ctgcccgggg cctccaggc gaggcctggg cctcacacac aggatctctc cgagcagcaa
601 ttccccattc ctctccccta ttgctggctc tgcagggtc tgatcaagcg gatccaagcc
661 atgattccca aggggtgcgt acgtgtggca gtggcccagg tgtgccgcgt ggtacctctg
721 gtggcgggcg gcatctgcca gtgcctggct gagegctact ccgtcatcct gctcgacacg
781 ctgctggggc gcatgtgtcc ccagctgggtc tgccgcctcg tcctccggtg tcccatggat
841 gacagcgtg gcccaaggtc gccgacagga gaatggctgc cgcgagactc tgagtgcac
901 ctctgcatgt ccgtgaccac ccaggccggg aacagcagcg agcaggccat accacaggca
961 atgctccagg cctgtgttgg ctccctggctg gacagggaaa agtgcaagca atttgtggag
1021 cagcacacgc ccagctgct gacctgggtg cccaggggct gggatgccca caccacctgc
1081 caggccctcg ggggtgtgtg gaccatgtcc agccctctcc agtgtatcca cagccccgac
1141 ctttga

```

Figure 3D. Human pulmonary surfactant-associated protein B, 18kD, amino acid sequence (SEQ ID NO: 4)

```

1 maeshllqwl llllptlcgp gtaawtsssl acagqpefwc qslegalqcr alghclqevw
61 ghvgaddlcq ecedivhiln kmakeaifqd tmrkfleqec nvlplklmp qcnqvlddyf
121 plvidyfqng idsngicmhl glcksrqpep eqepgmsdpl pkplrdplpd plldklvlpv
181 lpgalqarpg phtqdlseqq fpiplycwl cralikriqa mipkgalrva vaqvcrvvpl
241 vaggicqcla erysvilldt llgrmlpqlv crlvrcsmd dsagprspgt ewlprdsech
301 lcmsvttgag nsseqaipqa mlqacvgswl drekckqfve qhtpqlltlv prgwdahhcc
361 qalgvcgtms splqcihspd 1

```

Figure 3E. Human hypothetical protein FLJ40379 (ENSEMBL Accession Number ENSG00000173005), nucleic acid sequence (SEQ ID NO: 5)

```

1 ccatacatgc ttctggcagt gttgcacagc gttgcacagc ttggcggcct cctgggtcct
61 gcaccagaag cttggggccca gggcacactg gtcggtgccc agcagtgggg tcctggggcc
121 gtggcaggcc cccaccttct tgcacacagc cacggggtcc atcatgtcct tgagactctc
181 aatgagcacg ggctcgtact gggtgacgaa gtgcttgac tggatcatat agggcagcgg
241 caggatgctg cagccaccct tgaaggccac caggatgtct cgcttggtgc tcttgctctc
301 caagttgtgg gaggacaccg tgagcagcct cttgcaccca ttgcagaagc tgccctgggt
361 ctccgcgtcc cactctgggg acggcacgat ggcattaggca tcatggactg cccggggccg
421 cctccggttg ccacacagac ggatgaactt gcacaccttc tctgggggtga ttttggccac
481 aagctgcacc aaggaggggc tgtagggtgc caccaagatg atgcactcct tcgtgataga
541 ggaggcatt accgagcaca cgcgctccag ggcattgggtg atcatgagct cagagctgtt
601 ggacatgagc cagtgggtcca gcttctgcac caggttcacg cacacctcac aggtcacacc
661 ggccttcac tgcactctgc tctgtttcct tggcaacccc agctccaggg aggggacccc
721 gtccatggcc actacttgag tcaaacgggc aggtgccctt agctcctcac agaattcccc
781 cttcctgcag agctcctgcg gggggagaag cctcagtgtt tggtcagcag ggacaaaaaa
841 ctggaagagg tagttcttgc agaggacggc caggccaggc cccaaggact cacactgtct
901 ctggatgttc aagtcggcca aggtcaagtt ggaccggaca gcctcctgga gtcgggagac
961 ctgccgtaca cagtcttggc acagagctcc ttcaggcgcc tggcgggggt ggaaggtaag
1021 gggccattg gccatgaacg gagccacagc ctcaaagggt tcctctttgg agagtggcct
1081 cagggtggcc aggtgcctct gcagcggtc acagaggctg agcgctgtgc acacctgtgc
1141 cggggcactg tccggggccc cacggagcat gctcaggatg gccgaactgt gggcatccac
1201 catccacttg catccggctg aagactcctg gctggggagc cactcacagg tcttcacac
1261 caaagccagg atgtcagact ccgtggcgtc agggttcagc ccattgccag cggcggtgc
1321 tatgtcctgg catacgtcgc agggcagaga cttcgcggtg ggtttgttcc atacggcccc
1381 ttggcagtag cccacagccc cgcacctggc agctgtctgc agatcctgac accacaccgt
1441 ggagcccttt gcacactcct gggggcctga ggtggggctg gccctggtgg ccccaggag
1501 gctgggcagg aggagcaggg c

```

Figure 3F. Human hypothetical protein FLJ40379 (ENSEMBL Accession Number ENSP00000308224), amino acid sequence (SEQ ID NO: 6)

```

1 ALLLLPSLLG ATRASPTSGP QECAKGSTVW CQDLQTAARC GAVGYCQGA VWNKPTAKSLP
61 CDVCQDIAAA AGNGLNPDAT ESDILALVMK TCEWLPSQES SAGCKWMVDA HSSAILSMRL
121 GAPDSAPAQV CTALSLCEPL QRHLATLRPL SKEDTFEAVA PFMANGPLTF HPRQAPEGAL
181 CQDCVRQVSR LQEA VRSNLT LADLNIQEQC ESLGPGGLAVL CKNYLFQFFV PADQALRLLP
241 PQELCRKGGF CEELGAPARL TQVVAMDGVP SLELGLPRKQ SEMQMKAGVT CEVCMNVVQK
301 LDHWLMSNSS ELMITHALER VCSVMPASIT KECIILVD TY SPSLVQLVAK ITPEKVCKFI
361 RLCGNRRRAR AVHDAYAIVP SPEWDAENQG SFCNGCKRLL TVSSHNL ESK STKRDILVAF
421 KGGCSILPLP YMIQCKHFVT QYEPVLIESL KDMMDPVA VC KKVGACHGPR TPLLGT DQCA
481 LGPSFWCRSQ EAKLCNAVQ HCQKHVW

```

Figure 3G. Multiple Sequence Alignment (ClustalW 1.83)

```

PSAP Hs      -----MYALFLLASLLGAALAGFVLGLKECTRGSAVWCQNVKTASDCGAVKHCL
FLJ40379 Hs  -----ALLLLPSLLGATRASPTSGPQECAGSTVWCQDLQTAARCGAVGYCQ
SFTPb Hs     ----MAESHLLQWLLLLLPTLCGPGTAAWTTSSLACAQGPEFWCQSLEQALQCRALGHCL
Sap-r Dm     MERAGLLAVLALCCA-----FGVFAAATPLLGSCKTWGPSYWCNFSNSKECRATRHCI

PSAP Hs      QTVWNK---PTVKSPLPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLKPKNMSASC
FLJ40379 Hs  GAVWNK---PTAKSLPCDVQCQDIAAAGNGLNPDATESDILALVMKTCEWLPSQESSAGC
SFTPb Hs     QEVWGH---VGADDL-CQECEDIVHILNKMAKEAIFQDTMRKFLEQECNVLPKLLMPQC
Sap-r Dm     QTVWETQKVPVDTDSICTICKDMVTQARDQLKSNQTEELKEVFEGSCKLIPIKPIQKEC

PSAP Hs      KEIVDSYLPVILDIKIGEMS-RPGEVCSALNLCESL-----
FLJ40379 Hs  KWMVDAHSSAILSMRLRGAPDSAPAQVCTALSICEPL-----
SFTPb Hs     NQVLDDYFPLVIDYFQNQID--SNGICMHLGLCKSR-----
Sap-r Dm     IKVADDFLPELVEALASQMN--PDQVCSVAGLCNSARIDELYKNGIQAGLDGTVQNEDDS

PSAP Hs      -----QKHLAELNHQKQLESNKIPELDMTEVVAPFMANIPLLLYPQDGPGRSKPQ
FLJ40379 Hs  -----QRHLAT-----LRPLSKEDTFEAVAPFMANGPLTFHPRQAP-----
SFTPb Hs     -----QPEPEQEP-----GMSDPLPKPLRDPPLDPLLDKLVLVLPVLPALQ-----
Sap-r Dm     SEETELAMQPNQLSCGNCNLLSRLMHSKFAATDRDDMVETMLHMCGLSSFSDACANIVL

PSAP Hs      PKD-----NGD
FLJ40379 Hs  -----EGA
SFTPb Hs     -----
Sap-r Dm     TYFNDIYDHVSKHLTTDAVCHVSGVCASRYHQHEEEKQP--QEALVALDAG-----DDI

PSAP Hs      VCQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISQYSEIAIQMM
FLJ40379 Hs  LCQDCVRQVSRRLQEAVRSNLTAD---LNIQEQCESLGPGLAVLCKNYLFQFFVPADQAL
SFTPb Hs     -----ARPGPHTQDLSEQQFPPIPLP-----
Sap-r Dm     PCELCEQLVKHLRDVLVANTTETEFK--QVMEGFCKQSKGFKDECLSIVDQYYHVIYETL

PSAP Hs      MHMQ---PKEICALVGFC-----EVKEMPMQTL
FLJ40379 Hs  RLLP---PQELCRKGGFCEE-----LGAPARLTQV
SFTPb Hs     -----
Sap-r Dm     VSKLD--ANGACCMIGICQKNSAS--SMKDVPIMPLLPVIEPAQVKITIEKLEKHEKKQL

PSAP Hs      VPAK-VASKNVIPALELVEPIK---KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKE
FLJ40379 Hs  VAMD-----GVPSLELGLPRKQS---EMQMAGVTCEVCMNVVQKLDHWLMSNSSELM
SFTPb Hs     -----YCWLCRALIKRIQAMIPKG-----A
Sap-r Dm     GASEPKFSQQEILDMQLPIDHLMGAANPGALVEGGELCTLCEYMLHFIQETLATPSTDDE

PSAP Hs      ILDAFDKMC SKLPKSLSEECQEVVDITYGSSILSILLEEVSPVLVCSMLHLCSG-----
FLJ40379 Hs  ITHALERVCSVMPASITKECIIIVDTYSPSLVQLVA-KITPEKVCKFIRLCGNRRR-----
SFTPb Hs     LAVAVAQVCRVPLVAGGICQCLAERYSVILLDTLLGRMLPQLVLCRLVLRCSMDDS-----
Sap-r Dm     IKHTVENICAKLPSGVAGQCRNFVEMYGDAVIALLVQGLNPRDVCPLMQMCPKNLP-----

PSAP Hs      -----TRLPALTVHVTQPKDGGFCEVCKKLVGYLDRNLEKN
FLJ40379 Hs  -----ARAVHDAYAIVPSPWDENQGSFCNGCKRLLTVSSHNLISK
SFTPb Hs     -----AGPRSP-----TGEWLPDSECHLCMSVTTQAG-----N
Sap-r Dm     -----KKEDVEVFNPQPAS-----DEQDPPTCPLCLFAVEQAQMKIRDN

PSAP Hs      STKQEILAALEKGC SFLPDYPYQKQCDQFVAEYEPVLI EILVEVMDPSFVCLKIGACPSAH
FLJ40379 Hs  STKRDI LVA FKGGCSILPLPYMIQCKHFVTQYEPVLI ESLKDMMPVAVCKKVGACHGPR
SFTPb Hs     SSEQAIPQAMLQACVGSWLD-REKCKQFVEQHTPQLLTLVPRGWDAHTTCQALGVCCTMS
Sap-r Dm     KSKDNIKKVLNGLCSHLPNEIKEECVDFVNTYSNELIDMLITDFKQEI CVQLKLCPKTT

```

PSAP Hs KPLLGT-----EKCIW
FLJ40379 Hs TPLLGT-----DQCAL
SFTPb Hs SPLQ-----CIH
Sap-r Dm YALWDLRISLEDDVDGEDK-----SSSEEISFNDIESLEELPPQLAFDPGFTAAPNCLI

PSAP Hs GPSYWCQNTETAAQCNAVEHCKRHVWN-----
FLJ40379 Hs GPSFWCRSQEAAKLCNAVQHCKHVW-----
SFTPb Hs SPDL-----
Sap-r Dm CEELVKTLEKRMGKHPTRDSIKHILEESCDRMRKPMNTKCHKVIDKYGDKIADLLLKEMD

PSAP Hs -----
FLJ40379 Hs -----
SFTPb Hs -----
Sap-r Dm PKLICTELGMCILAD--LDDLEVDEALKYDVIALPRQDNKLS-----SSIKEPPTCVLCE

PSAP Hs -----
FLJ40379 Hs -----
SFTPb Hs -----
Sap-r Dm FIMTKLDADLKNKTEQDDIKRAIEAVCNRLPATVRKQCDTFVDGYASAVLKLLSDVPPKQ

PSAP Hs -----
FLJ40379 Hs -----
SFTPb Hs -----
Sap-r Dm VCQKLQLCFSVAVTD-----EVLECGVCHGVTQALLPFLREKKDNVSEVTALQMTSVGCE

PSAP Hs -----
FLJ40379 Hs -----
SFTPb Hs -----
Sap-r Dm NLPKYYKICSEMISYIGSSIKNLAKRPYIDQSHICAEIGKCFESEKSSLAFARISA

Figure 4. Expression of the Sap-r homolog Psap in mammalian tissues

Figure 4A. Quantitative analysis of Psap expression in wildtype mouse tissues

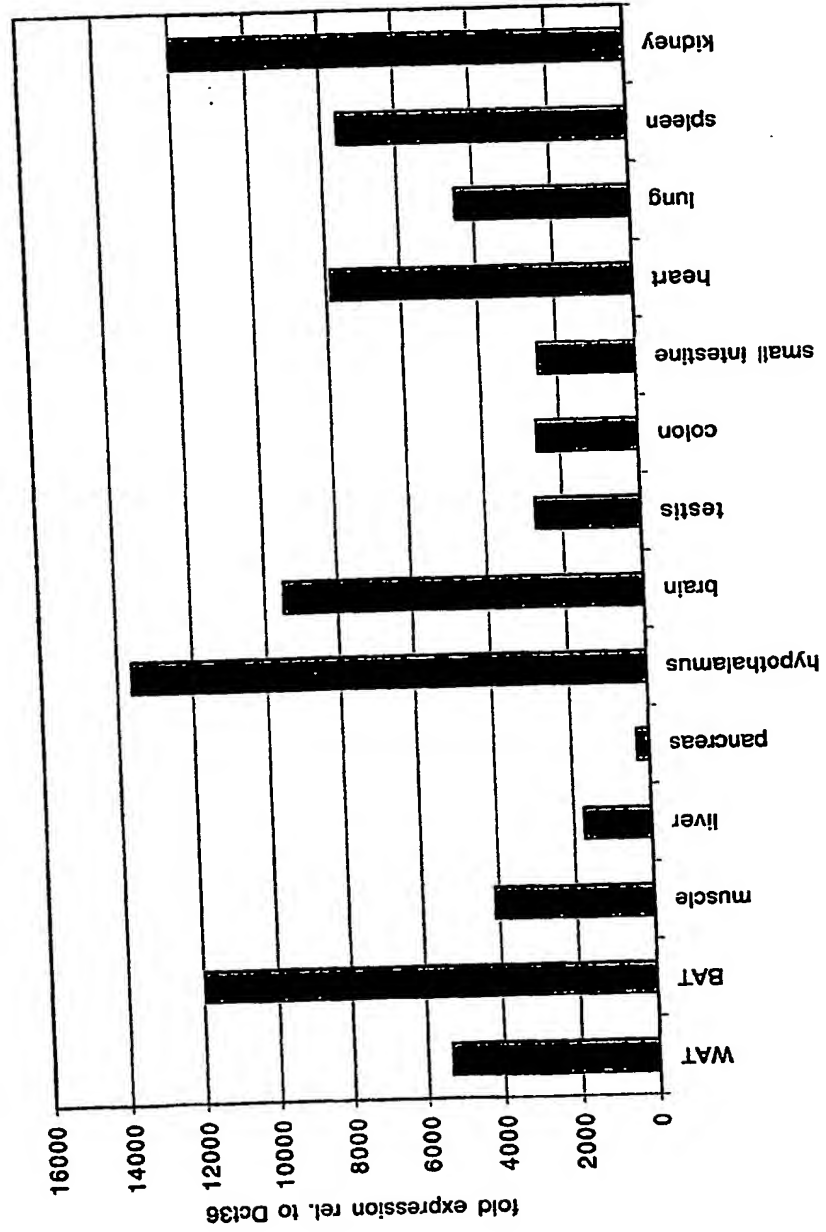


Figure 4B. Quantitative analysis of Psap expression in different mouse models

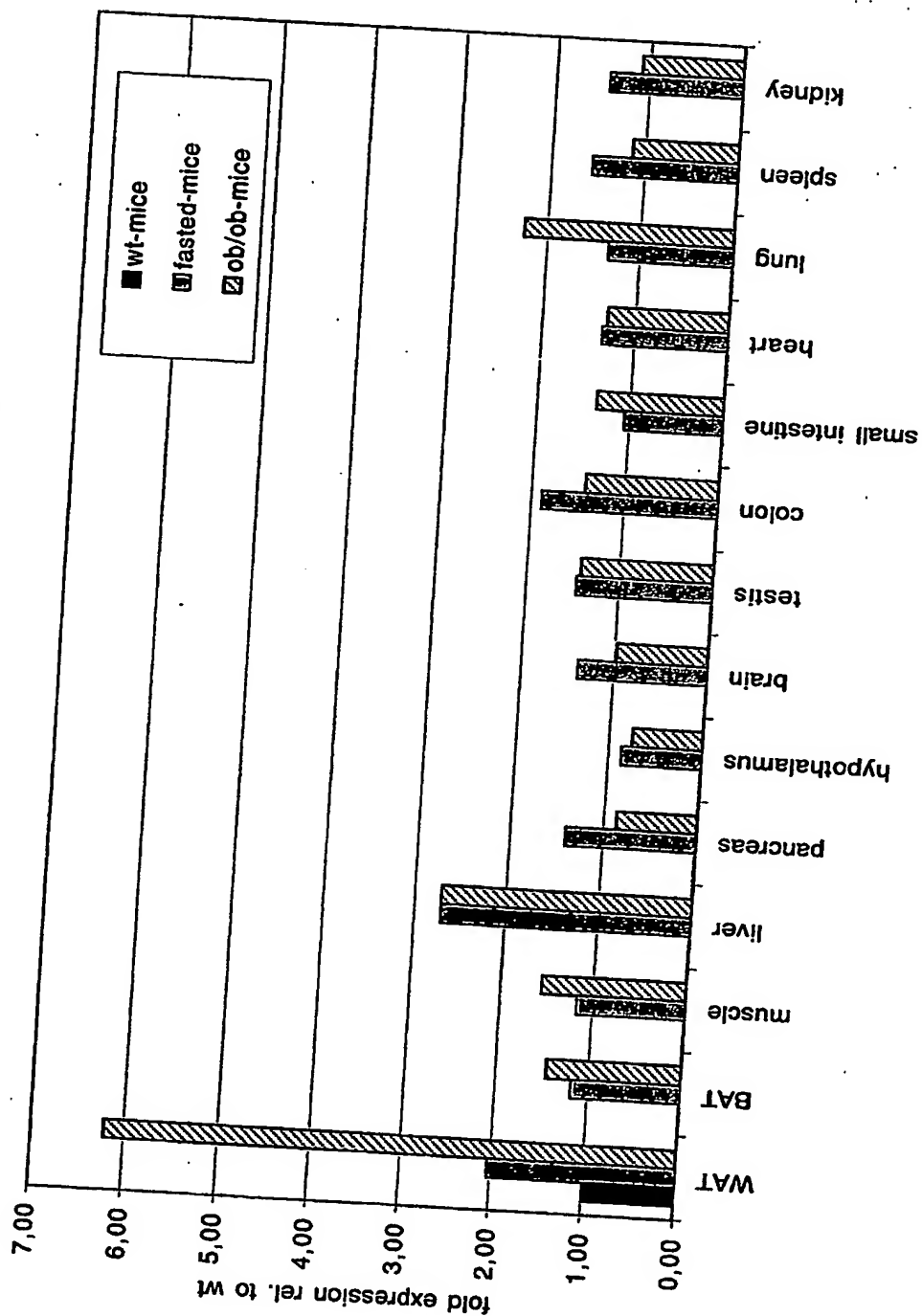


Figure 4C. Quantitative analysis of Psap expression in mice fed with a high fat diet

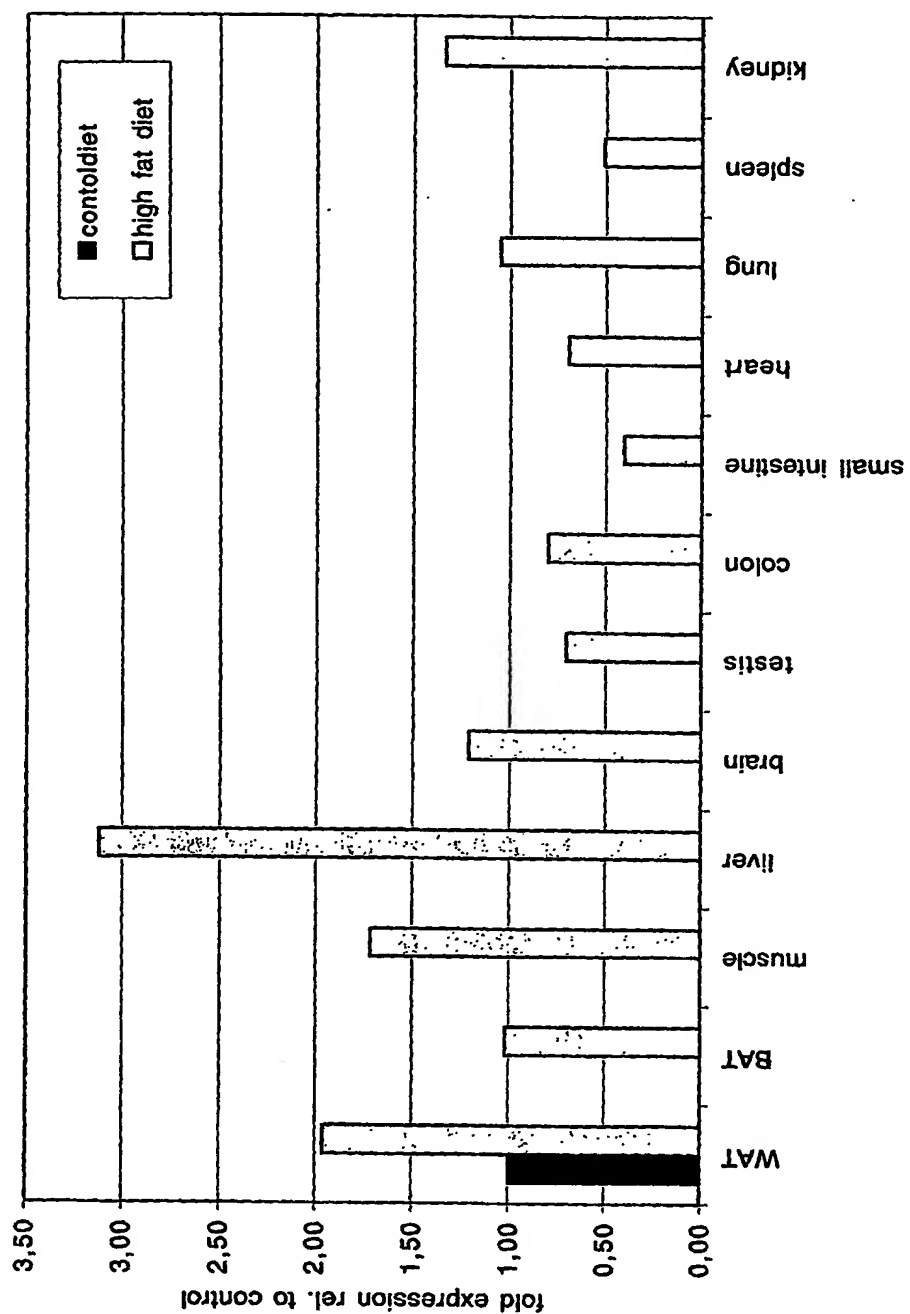
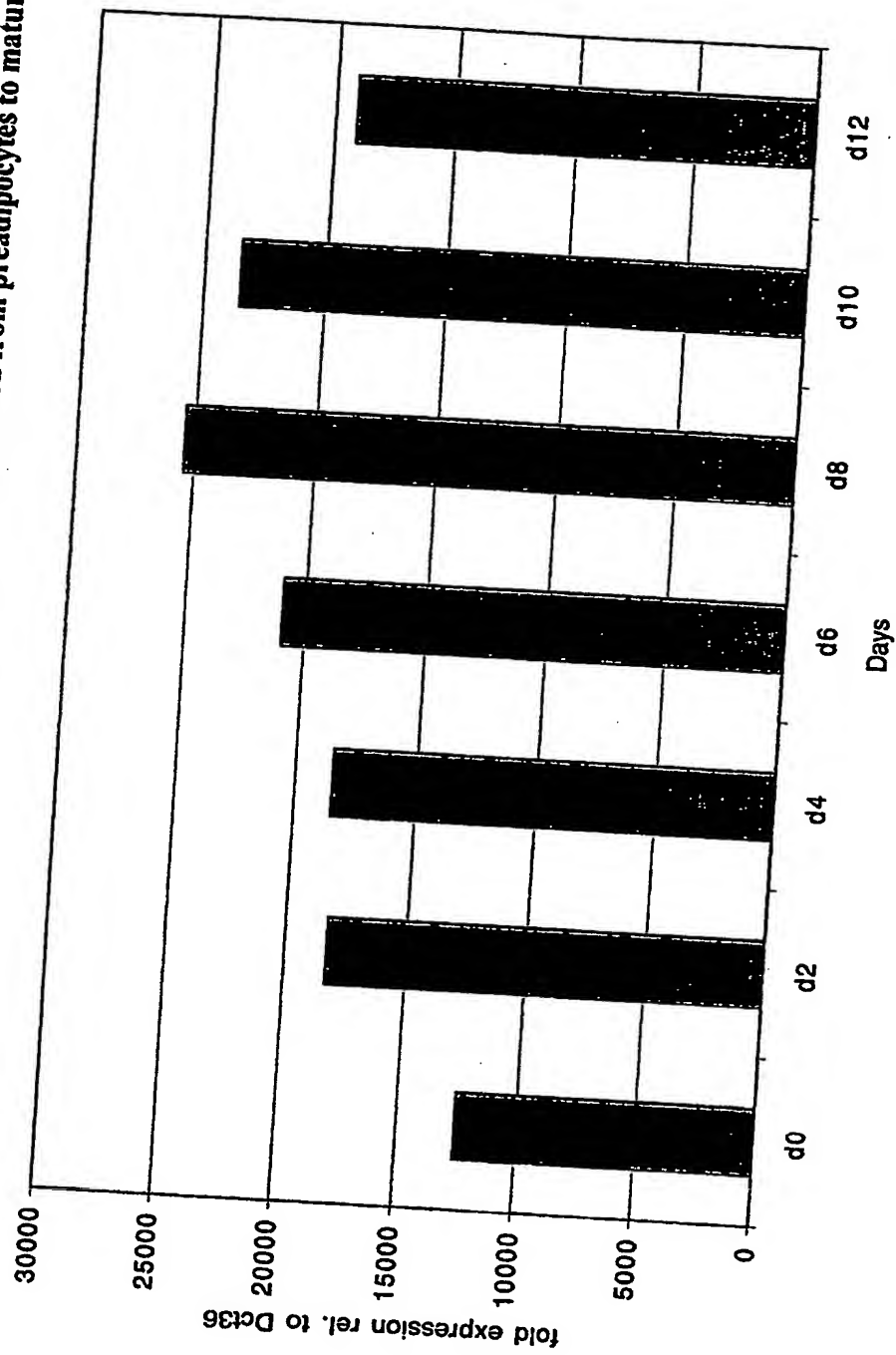


Figure 4D. Quantitative analysis of Psap expression 3T3-L1 cells differentiated from preadipocytes to mature adipocytes



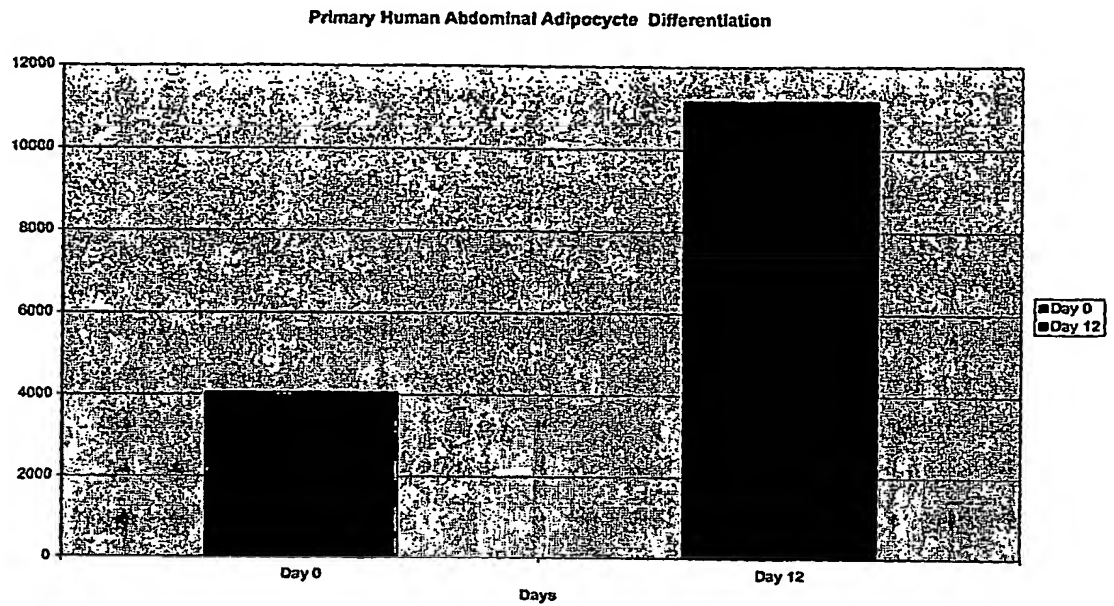


FIGURE 5. Quantitative analysis of prosaposin expression in human abdominal adipocyte cells during the differentiation from preadipocytes to mature adipocytes

PCT/EP2004/003244



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.